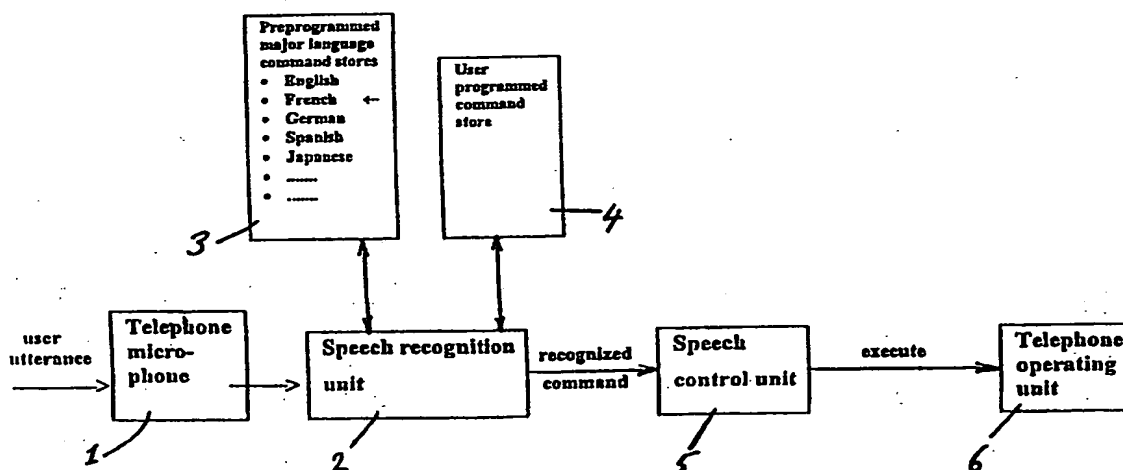




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : <b>G10L 15/02</b>	<b>A1</b>	(11) International Publication Number: <b>WO 00/22609</b> (43) International Publication Date: <b>20 April 2000 (20.04.00)</b>
<p>(21) International Application Number: <b>PCT/SE99/01833</b></p> <p>(22) International Filing Date: <b>11 October 1999 (11.10.99)</b></p> <p>(30) Priority Data: 09/170,003      13 October 1998 (13.10.98)      US</p> <p>(71) Applicant: <b>TELEFONAKTIEBOLAGET LM ERICSSON</b> (publ) [SE/SE]; S-126 25 Stockholm (SE).</p> <p>(72) Inventors: <b>HERMANSSON, Hans</b>; Godsägarvägen 21, S-121 63 Johanneshov (SE). <b>DOBLER, Stefan</b>; Tennenbachweg 2a, D-91077 Neunkirchen am Brand (DE). <b>HIRSCH, Hans-Günter</b>; Bühlweg 1A, D-90562 Heroldsberg (DE). <b>BRUHN, Stefan</b>; Thumenberger Weg 38, D-90491 Nuremberg (DE).</p> <p>(74) Agents: <b>PLATT, T. et al.</b>; Albihns Patentbyrå Stockholm AB, P.O. Box 5581, S-114 85 Stockholm (SE).</p>		<p>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p><b>Published</b> With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</p>

(54) Title: **SPEECH RECOGNITION AND CONTROL SYSTEM AND TELEPHONE**

## (57) Abstract

A speech recognition and control system suitable for mobile telephones has, for each of four or five major languages, a preprogrammed store containing many variations of a set of telephone operating commands. The user can manually select one of these four or five major languages, and the selected preprogrammed language store will be consulted when the user utters a word into the telephone. A match (recognition) prompts execution of the desired telephone function. The user can replace each of the preprogrammed commands with his own user-chosen and -spoken commands to create his own set of commands specific to his own native language/dialect and/or pronunciation. The user can also add additional user-dependent commands and a personal user-defined telephone directory.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

**Title: Speech Recognition and Control System and Telephone****TECNICAL FIELD**

5 The invention relates to speech recognition and control systems, particularly to those which can be utilized by a user to control a telephone by spoken commands or by a combination of spoken and manual commands.

**BACKGROUND OF THE INVENTION**

10

An apparatus such as a telephone, which is to be controlled by spoken commands must have a system for recognizing the speech received by the microphone. There are in essence two different categories of speech recognition systems, speaker independent systems and speaker dependent systems.

15

A speaker dependent speech recognition system is alterable to accommodate the individual user of the system, for example the owner or owners of a mobile telephone, operated in response to spoken commands. Speaker dependent speech recognition systems can recognize the command words as pronounced by the individual user. It is also possible for the individual user to use words in his own native language or user-created language to trigger the various functions of the apparatus, for example a mobile telephone. To do this, however, it is first necessary for the user to train the recognition system by going through a long and cumbersome programming routine, in which *each* command is repeated several times by the user.

20 This must be done before the apparatus can be used at all. Such a system does not permit anyone else to use the system, without going through the same cumbersome initial procedure.

25

30

A speaker independent system is essentially a system which can recognize spoken words in the vocabulary regardless of variations in the speaker's speech depending

on sex, age and accent. A large number of different speakers must be sampled in order to provide a broad spectrum of system recognizable pronunciations of a particular word. A speaker independent speech recognition system has the advantage that it can be used immediately without any initial "training" of the system to recognize the words in the vocabulary. Ideally, a speaker independent speech recognition system would be so broad as to recognize all different possible pronunciations and accents as well as having a separate language mode for each different language in the world. This ideal system is, however, hardly practical, even if one only tries to cover languages spoken in Europe. The broader the recognition base and the more languages are included, the more laborious, extensive and expensive the creation of the speaker independent system will be. And even more work will be involved as further functions and modifications are to be included in the speech recognition and control system.

## DESCRIPTION OF RELATED ART

Many different speech recognition systems of the above types have been developed. One such system is described in WO96/13827 (PCT/GB95/02563) to Ringland et al. This known system is based on the recognition of individual phonemes (subwords) which are then combined to form commands which control the various functions of the apparatus. Instead of recognizing complete words, which theoretically can be infinite in number, this known system recognizes phonemes which are the building blocks of words and are finite in number. After positively identifying a phoneme, a processor combines it with other positively identified adjacent phonemes to create a word or pose. This system is much more economical as regards storage capacity. During use of this system, when a specific phoneme is recognized, by comparison with a predefined store of standard phonemes, the actual user utterance of this phoneme, with his particular inflection and accent, is stored in another parallel memory, thereby improving future recognition of this phoneme when uttered by this particular speaker.

This known system does not, however, allow the user to substitute another language, for example if a Swedish user wishes to say "slå" instead of "dial". And also it does not permit major deviations from the standard pronunciation of a specific phoneme.

## 5 SUMMARY OF THE INVENTION

The present invention overcomes these and other shortcomings with a speech recognition and control system and a telephone incorporating this system. The speech recognition and control system, which is included in the telephone according to the invention, is preprogrammed to recognize at least one speaker independent set of audible commands. Several different sets of basic speaker independent audible commands, one set for each of several major languages such as for example English, French, German, Spanish, Japanese etc. could be produced by the manufacturer of the system and/or telephones and be preprogrammed into the system. If the system is incorporated in a mobile telephone for example, the appropriate language can be selected manually via the menu language selection function. Thereafter, the user can utter one of the limited number of preprogrammed audible commands to the system in the selected major language.

20 This does not require any initialization or system programming on the part of the user. For each of the words in the preprogrammed set of commands, a broad spectrum of different standard pronunciations are recognizable by the system in the manually selected major language.

25 The user is then able, as he is using the system, to add user-dependent commands. perhaps relating to personal telephone directory entries. The system according to the invention is also configured so that each of the preprogrammed commands may be replaced by a user-specific utterance. This may be the user's dialect pronunciation of one of the commands in the selected preprogrammed set of commands or it may be a corresponding command in the user's native language which is not one of the major

30

preprogrammed languages. A Swedish user may replace the English command "dial" with the Swedish command "slå" so the system will be adaptable to virtually any language.

- 5 Furthermore, the replacement feature permits the user to replace the standard commands with any code or imaginary commands of the user's choosing.

The system according to the invention is also provided with a user-actuated function to return the system, temporarily or permanently, to the original preprogrammed sets  
10 of recognizable commands.

The speech recognition and command system according to the invention thus provides the advantages of a user-independent ready-to-use system together with the versatility and customization of user-specific and user-defined systems, without the  
15 disadvantages thereof.

## DESCRIPTION OF THE DRAWING

The accompanying figure illustrates schematically one embodiment of the speech  
20 recognition and command system incorporated in this particular example in a mobile telephone.

## DETAILED DESCRIPTION

25 The figure shows a block diagram of a speech recognition and control system incorporated in a telephone in accordance with the present invention. The speech recognition unit 2 and the speech control unit 5 are incorporated in a telephone. Preprogrammed command stores 3 in several major languages such as English, French, German, Spanish, Japanese, etc. are coupled to the speech recognition unit.  
30 In each of these language stores a set of command words is stored with many

different standard variants, to cover a broad spectrum of different pronunciations, tonal levels etc., of each command to be recognized.

The user first selects the major language which he wishes to use initially. This  
5 can be done, for example, by selecting manually a desired language from a menu  
presented in the telephone display window. In this case the arrow in the box 3 indicates that French has been selected as the initial major language for operating the  
telephone by voice commands. The user can then utter commands in standard  
French in the telephone transmitter/microphone. The user may say the command  
10 "compose" (dial) and the speech recognition unit will check with the command store  
for French, which has already been manually selected, to see if the audio signal  
generated by the user saying "compose" matches any of the variants of this command stored in the command store for French. If there is a match, then the speech  
recognition unit 2 will send a signal that it has recognized the command for "dial" to  
15 the speech control unit 5. The speech control unit will in turn send an "execute"  
signal to the operating unit 6 of the telephone to perform the operation "dial".

A user-programmed command store 4 is arranged in parallel with the preprogram-  
med major language command stores 3. The user may enter utterances in the user-  
20 programmed command store to replace the commands in the selected major  
language command store. With the telephone in "replace" mode, the user can give  
the command "compose", having chosen French as his initial operating language  
and then give the word for dial in his own native language, saying "slå" if his native  
language is Swedish. He can thus replace all of the commands in the standard set  
25 of preprogrammed commands with commands in his own particular language,  
dialect or pronunciation. He can even enter secret code words if he wishes. The  
system is also provided with an override function to ignore the commands entered in  
the user-programmed command store and use the selected major language command  
store instead. This will enable another user to be able to use the speech recognition  
30 and control functions of the telephone.

The user can also enter his own set of recognizable additional commands to compile his own personal telephone directory for example. Each stored number may be coupled to a user-uttered name, and this name can follow the command "dial/compose/slå" to dial the number of the desired person. The numbers in  
5 the personal telephone directory may be entered manually, automatically or by audio recognition.

It is also envisioned that the user will be able to add additional commands to the user-programmed command store. For example, dormant functions may be activated  
10 and be controlled by a user-defined voice command. One simple example would be the display of the remaining battery charge by uttering any selected user-defined command. The current time in any selected time zone might also be displayed by uttering user-defined commands.

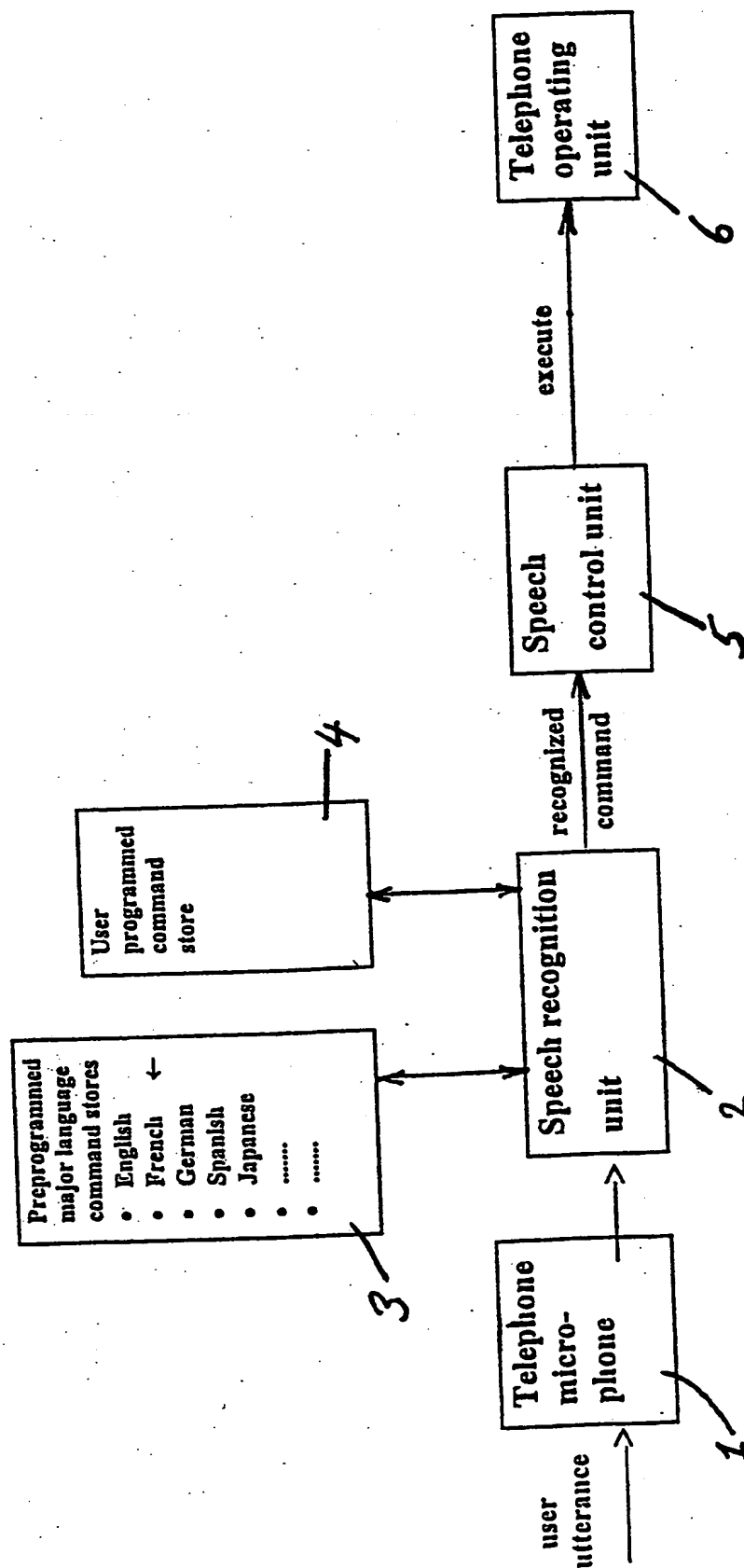


## CLAIMS

1. A speech recognition and control system comprising:  
a microphone for receiving audible commands spoken by a user and generating  
5 electrical audio signals,  
a processor including a recognizer for recognizing received electrical audio signal  
patterns and means for generating command signals in response thereto, said  
processor being coupled to said microphone to receive its generated electrical audio  
signals,  
10 wherein said processor is preprogrammed to recognize at least one speaker  
independent set of audible commands, some or all of said audible commands being  
individually replaceable by the user with user-chosen and -spoken audible  
commands which are user-specific.
- 15 2. Speech recognition and control system as defined in Claim 1, wherein said  
processor is preprogrammed to recognize a plurality of speaker independent sets of  
audible commands in a plurality of different languages.
- 20 3. Speech recognition and control system as defined in Claim 1, wherein the  
processor is disposed to incorporate additional user-specific spoken audible  
commands entered by the user.
4. Telephone, comprising a speech recognition and control system as defined in  
Claim 1, 2 or 3, wherein said microphone is a telephone transmitter and said  
25 processor generates command signals to operate the telephone.
5. Telephone as defined in Claim 4, wherein a preprogrammed speaker independent  
set of audible commands in a desired language can be employed by manually  
selecting a desired language mode on the telephone.

6. Telephone as defined in Claim 4, wherein the telephone is a mobile telephone.

7. Telephone as defined in Claim 4, wherein the additional user-specific spoken audible commands comprise a list of names with related telephone numbers.



Figure

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/SE 99/01833

## A. CLASSIFICATION OF SUBJECT MATTER

IPC7: G10L 15/02

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: G10L, H04M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 9618258 A2 (VOICE CONTROL SYSTEMS, INC.), 13 June 1996 (13.06.96), page 2, line 1 - page 3, line 2; page 7, line 15 - page 8, line 3 --	1-7
X	WO 9613827 A1 (BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY), 9 May 1996 (09.05.96), page 4, line 21 - line 34; page 7, line 13 - line 31 --	1-7
A	WO 9502879 A1 (BORDEAUX, THEODORE, AUSTIN), 26 January 1995 (26.01.95), abstract -----	2,5



Further documents are listed in the continuation of Box C.



See patent family annex.

## \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "B" earlier document but published on or after the international filing date
- "I" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

8 March 2000

Date of mailing of the international search report

2000-03-08

Name and mailing address of the ISA /

Swedish Patent Office  
Box 5055, S-102 42 STOCKHOLM  
Facsimile No. +46 8 666 02 86

Authorized officer

Jack Hedlund/ELY  
Telephone No. +46 8 782 25 00

# INTERNATIONAL SEARCH REPORT

Information on patent family members

02/12/99

International application No.

PCT/SE 99/01833

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9618258 A2	13/06/96	AU 705525 B	27/05/99
		AU 4739596 A	26/06/96
		CA 2206870 A	13/06/96
		EP 0795246 A	17/09/97
		US 5881134 A	09/03/99
WO 9613827 A1	09/05/96	AU 707355 B	08/07/99
		AU 3751695 A	23/05/96
		EP 0789901 A	20/08/97
		FI 971822 A	29/04/97
		JP 10507536 T	21/07/98
		NO 972026 A	30/04/97
		NZ 294659 A	28/01/99
WO 9502879 A1	26/01/95	AU 682380 B	02/10/97
		AU 7328294 A	13/02/95
		CA 2167200 A	26/01/95
		EP 0708958 A	01/05/96
		JP 9500223 T	07/01/97
		US 5758023 A	26/05/98